GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:32 ; Search time 170.72 Seconds
                                          (without alignments)
                                          248.582 Million cell updates/sec

Title: US-09-331-631A-5
Perfect score: 3326
Sequence: 1 QCMQLETSGQMRRCVSQCDK.........SPRSTKQQQPLVSILDFVGF 625

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_66:*
               1: pir1:*
               2: pir2:*
               3: pir3:*
               4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1163	35.0	810	2	T44430	protein PV100 (imp
2	1112	32.3	566	2	S21825	vicilin-like stora
3	1074.5	32.3	588	1	S22477	vicilin precursor
4	1063.5	33.0	605	2	FWPMVA	vicilin A precurso
5	1018.5	30.6	605	2	FWCNAB	alpha-globulin B p
6	903	30.5	509	2	S06398	globulin-1 - maize
7	888	27.1	573	2	S08059	alpha-globulin typ
8	864.5	26.7	582	2	A53234	alpha-globulin typ
9	859.5	26.0	637	2	B53234	vicilin-like stora
10	858.5	25.8	605	1	S35221	globulin-1S, GLB1S
11	854.5	25.8	605	2	FWSYBA	globulin Beg1 prec
12	833	25.7	605	1	FWSYBA	beta-conglycinin a
13	794	25.0	639	2	S20007	beta-conglycinin a
14	775.5	23.9	571	2	B24810	beta-conglycinin a
15	759.5	25.0	524	2	S00566	convicilin precurs
16	759.5	22.8	483	2	JQ1730	62K sucrose-bindin
17	757.5	22.8	433	2	T06459	62K sucrose-bindin
18	749	22.8	438	2	S00567	vicilin precursor
19	740	22.5	439	1	S05505	vicilin, 47K - gar
20	722	22.2	448	2	S35757	vicilin, 47K - gar
21	717.5	21.7	450	2	FWSYCB	beta-conglycinin b
22	713.5	21.6	463	2	S24756	vicilin-like stora
23	712.5	21.5	463	2	S15675	globulin-2 precurs
24	712.5	21.4	463	2	A27288	vicilin precursor
25	692.5	20.8	410	1	S05309	vicilin precursor
26	685.5	20.6	445	2	FWPMWB	vicilin B precurso
27	563	16.9	445	2	JQ2264	canavalin - Jack b
28	560	16.8	436	2	S00281	canavalin - sword
29	558.5	16.8	430	2	A23498	phaseolin type alp
27	560	16.8	430	2	S10056	phaseolin type alp
28	558.5	16.8	414	2	S60190	alpha-phaseolin pr
29	552	16.5	421	2	A24810	vicilin - Zamia fu

(additional description column entries)
vicilin-like stora
vicilin A precurso
globulin-1 - maize
convicilin (clone
probable major pro
convicillin precurs
globulin1 - maize
beta-conglycinin a
7S storage protein
hypothetical prote
phaseolin - hum
trichohyalin - kidney
globulin 2 precurs
prunin 1 precursor
hypothetical prote
trichohyalin - rab

ALIGNMENTS

RESULT 1
T44430
protein PV100 [imported] - winter squash
C;Species: Cucurbita maxima (winter squash)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T44430
R;Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
J. Biol. Chem. 274, 2563-2570, 1999
A;Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a sin
A;Reference number: Z22767; MUID:99107919
A;Accession: T44430
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-810 <YAM>
A;Cross-references: EMBL:AB019195; NID:g3808061; PIDN:BAA34056.1; PID:g3808062

Query Match   35.0%; Score 1163; DB 2; Length 810;
Best Local Similarity   36.8%; Pred. No. 6.1e-66;
Matches 271; Conservative 133; Mismatches 192; Indels 140; Gaps 22;

```
Qy     10 QMRRCVSQCDKRFEE--------------------------------DIDWSKY----------------- 31
          |:|:|                                          |:|:
Db     94 QQRKCEQVCEERLREREQGRGEDVDEVERRDPEWEREQORRHEHEREERRRRERERER 153

Qy     32 ------------DNQEDPQTECQCQCQRCRCRQQESDPRQQYCQRCKEICEEEEYNRQRDPQQQY 85
                       ::|| || :||| |||||||  :|  |||||:|  :|
Db    154 GRGRRDENERDPKREQEERQR--REQERRRREQEQRERRGRDEEDDEN-QRDPDWRR 210

Qy     86 EQCQKRCQRRETEPRHMQICQQRCQERRY-----EKE-----KRKQQKRYEEQQREDE 132
          |:|| |||   |||||  | |||||||      |||     |||||||||||||||
Db    211 EQ----ERREQERRRREQEQERRERQRGRGRDDEDENQRDPDWREQKRREQEQERRER 265

Qy    133 EKYEERMKEGDNKRDP------QOREYEDCRRHCEQQ--------------------EP--R 166
          |:| | |||| ||:|        |:||||| | |                      |:
Db    266 ERRGGRDDEDENQRDPDWRREQERREQERRRREQEOERREHRGRGRDVEDENQRDPDWR 325

Qy    167 LQYCQCORRCEQ------------OROHGRGG------------DLMNPQRGGSGRYEE----GEEK 205
          ::|  |   |                ||||||              | ||  ||    || :
Db    326 REQERRRKEEQEREWEREHGRGRREQEEQRSREDERRRHEROHGGRSRVNQVAIRRTEQE 385

Qy    206 QSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTH 265
          |:|||:||||||:||||||||||:|||||||||:|||||||:|||||||:||:|:|||
Db    386 QSNNPYYFQEQRFQSRYRSDEGHWRVLRFSERSELLKGIKNQRLALLEARPHTFIVPHH 445

Qy    266 LDADALLLVIGGRGALKMIHRDNR------ESYNLECGDVIRIPAGTTFYLINRDNNERLHI 321
          |||:|||| |||||||:||:||:|      |:||:|:|||:||||||||||:|||||||||
Db    446 LDAECVLLVVRGRATITTVVQEKRETRKESYNVESGDVWTIPAGTTLYLYANQE-NEDLQI 504

Qy    322 AKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVLGQQRE--GV 379
          |:|  |:|:  ||:  ||   |||||||  |:||||||||||:|||||:|| || :||
Db    505 VKLVQPVNNPGEPKDYLSAGGES-QAYSVFSNDVLEAALNIPRDKLERIFKORRERGGK 563
```

```
Qy    380  IIRASQEQIRELTRDDSESRRWHIRRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR 439
           ||:|  |||||||||||||||||||||||||||| |:| ||| |||||:||| |||| |:|
Db    564  IVRASQOLRALSQ----RATSVRGSRGVRAPIKLESQPVYNNQYGQMFEACPDEFP 618

Qy    440  QLQDMDVSVFIANITQGSMMGGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGK- 498
           |: ||||||||||||||| |||||||||||||| ||||||| |||||| |||||||||||
Db    619  QLRRTDVATSVVDIKQGGMMVPHFNSRATWVVFVSEGAGSFEMACPHIQSSSQWQRGREE 678

Qy    499  ---RHEEEEE-----VHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLFAFGINAQ 549
             :|||||||     |||:|||||||||||||||||||||||||||||||||||||||
Db    679  ERHWREEEEEERSGRFERVAGRLSEGGVLVIPAGHPIAIMASPNENLRLVGFGINAE 738

Qy    550  NNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELLFNSQDESIFFPGPRQHOQQSPRS 609
           |||||||||||||||||||||||||||||||||||||||||||||||| |||  ||||||
Db    739  NNHRNFLAGRE-NIMNELDREAKELAFNVEGKQADEIFRSQRESFFTEGP---EGGRRR 793

Qy    610  TKQQOPLVSILDFVGF 625
           ||:|   |  | |:
Db    794  STERSPLLSILKLAGY 809
```

```
Query Match      33.4%;   Score 1112;  DB 2;   Length 566;
Best Local Similarity   40.6%;  Pred. No. 6.6e-63;
Matches  234;  Conservative  106;  Mismatches  179;  Indels  58;  Gaps  13;
```

```
Qy     68  ICEEEEEYNR------QRDEQQOQYEQCQKRCQRRETEPRHMQICQQQRCERRYEKEKRQOQKRY 124
             ||||  || |       | | |||||||||||||||||||||||||||| | |||||||||
Db     22  LCSGVSAYGRKQYERDPROQYEQCQQRRCESEATEERREQEQCEQRCER--------EY 70

Qy    125  EEQOQREDEEKYEERMKEGDNKRDPQQREYEDCRRHCEHQQE--PRLQYCQQRCEQQOROH 182
             ||||| | ||| |  || ||||||  | | | |  |||||  |||||| | ||| |
Db     71  KEQOROQEEEL------------ORQYQQCQGRCQEQOQRCWEQYKEQ 116

Qy    183  GRGGDLMNPQRGGSGRYEEGGEEKQSDNPYYFDE-RSLSTRRFRTEEGHISVLENFYGRSKL 241
            | |||| || | |||| ||| || | |||||| ||:|||| | |||||:||||||||||
Db    117  ER-GEHENYHNHKRSEEEGQORNNPYFPFKRSFQTRFRDEEGNFKILQRFAENSPP 175

Qy    242  LRALKNYRLVLLEANPNAFVLPTHLDADALLVIGGRGALKMIHRDNRESYNLECGDVIR 301
             | | ||| | | |:   |  |  || | || || |||||:||||:|||||||||||||
Db    176  LKGINDYRLAMEEANPNTFILPHHCDAEAIYFVTNGKGTIFEVTHENKESYNVQRGTVVS 235

Qy    302  IPAGTTFYLINRDNNERLHIAKFLQTISTPGYKEFFPAGQNPEPYLSTFSKEILEAAL 361
            |||||||||||||||||||| |||| |   || | |||| ||| | | ||  | | |
Db    236  VPAGSTVVVSQDNQEKLTIAVLALPVNSPGKYELFFPAGNNKPESYIGAFSYEVLETVF 295

Qy    362  NTQTERLRGVLGOQR------EGVIIRASQEIRELTRDDSESRRWHIRRGGESSRGP 413
            ||:|  |||| | | ||      :|||||||||||||||||||||||||||||||||||
Db    296  HLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNNERLHIAKF 324
```

```
Query Match      32.3%;   Score 1074.5;  DB 1;   Length 588;
Best Local Similarity   38.3%;  Pred. No. 1.6e-60;
Matches  229;  Conservative  115;  Mismatches  181;  Indels  75;  Gaps  13;
```

```
Db    296  NTQREKLEDILEEQRGQKRQQCQOQGMFRKAKPEQIRALSQQATSPR----HRGGE--RLA 349

Qy    414  YNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDVSVFIANITQGSMMGGPFFNTRSTKVVV 473
           |||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db    350  INLLSQSPVYSNQNGRFFEACPEDFSQFQNMDVAVSAFKLNQGAIFVPHYNSKATFVVFV 409

Qy    474  ASGEADVEMACPHLSGRHGGRGGGK--RHEEEEEV----HYEQVRARLSKREAIVVLA 525
           ||||| ||||||  |||| ||| |   |||||||     ||||||||||| | | |||
Db    410  TDGYGYAQMACPHLSRQSQGSQSGRQDRREQEEEEETFGEFQQVKAPLSPGDVFVAPA 469

Qy    526  GHPVVFVSSGNENLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEE 585
           ||| ||||||||||||||||||||||:||||||||||||||||||||||||:|||||
Db    470  GHAVTFFASKDQPLNAVARGLNAQNNQRIFLAGK-KNLVRQMDSEAKELSFGVPSKLVDN 528

Qy    586  LFNSQDESIFFPGPRQHOQQSPRSTKQQOPLVSILDF 622
           ||||||||||||| |  |||||||   || | ||||
Db    529  IFNNPDESYFMSFSQQORGDER---RGNPLASILDF 562
```

```
RESULT 4

S06398
alpha-globulin type A precursor - upland cotton
N;Alternate names: seed storage protein
C;Species: Gossypium hirsutum (upland cotton)
A;Accession: S06398
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
R;Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
  Plant Mol. Biol. 9, 533-546, 1987
A;Reference number: S06398
A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se
A;Accession: S06398
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-605 <CHL>
A;Superfamily: glycinin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-605/Product: alpha-globulin type A #status predicted <MAT>

Query Match     32.0%;  Score 1063.5;  DB 2;  Length 605;
Best Local Similarity     38.0%;  Pred. No. 8.2e-60;
Matches 232;  Conservative 121;  Mismatches 195;  Indels 63;  Gaps 16;

Qy   33 NQEDPQTECQOCCQRRCRQOQESDPQOQYCQRCKEICEEEEYNRQRDPQOQYEQCQKRC 92
Db   33 SEDDPQORYEDCRKRCQLETRCQTEQDKCEDRSSETQLKEEQORDGE-DPQORYQDCRQHC 91

Qy   93 QRRETEPR-HMQICQQRCERRYEKEKRKQQKRYEEQQOREDEEKYEERMKEGDNKRDPQOR 151
Db   92 QOEERRLRPH----CEOSCREQYEK----------QQQQOPD 120

Qy  152 EYEDCRRHCEQQE---PRLQYCQCORCEQEQEQORQHGRGGDLMNPQR-GGSGRYEEGEEKQSD 208
Db  121 QFEKCQQROCQWQFORPERKQOQCVKECRRQOYQEDPWKGERENKWREEEEEEESDEGEQOQRN 180

Qy  209 NPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDA 268
Db  181 NPYFHRRSFQERFREEHGNFRVLQREADKHHLLRGINEFRIAIEEANPNTFVLPHHCDA 240

Qy  269 DAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPACTTFYLINRDNNERLHIAKFLQTI 328
Db  241 EKIYVVTNGRGTVTFVTHENKESYINVPGVVVRIPAGSTVVLANQDNREKLITIAVLHRPV 300

Qy  329 STPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVLG------QQREGVII 381
Db  301 NNPGQFOKFFPAGQNPQSYLRFSREILEAVFNTRSEQLDELPGGRQSHRRQQQMFR 360

Qy  382 RASQEQIRELTRDDSESRRWHIRRGGESSRG-PYNLFNKRPLYSNKYGQAYEVKPEDY-R 439
Db  361 KASQEQIRALSQGATSPR------GKGSEGYAFNLLSQTPRYSNQNGRYEACPRNFQQ 413

Qy  440 QLQDMDVSFIANITQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR 499
Db  414 QLREVDSSVVAFEINKGSIFVPHYNSKATFVVLVTEGNGHVEMVCPHLS-RQSSDWSSRE 472

Qy  500 HEEEEEV------------HYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLLFAFGI-NAQNN 551
Db  473 EEEQQEQEVERRSSGQYKRVRAQLSTGNLFVVPAGHPVTFVASQNEDIGLLGFGLYNGQDN 532

Qy  552 HENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFNSQDESIFFPGPRQHQQQSPRSTK 611
Db  533 KRIFVAGKTNNV-ROWDRQAKELAFGVESRLVDEIFNNNPQESYFVSGRDRRGFDER-RG 590

Qy  612 QQOPLVSILDF 622
Db  591 SNNPLSPFLDF 601
```

```
RESULT 5

S08059
alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)
N;Alternate names: seed storage protein
C;Species: Gossypium hirsutum (upland cotton)
A;Accession: S08059
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
R;Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
  Plant Mol. Biol. 9, 533-546, 1987
A;Reference number: S06398
A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
A;Accession: S08059
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-509 <CHL>
C;Superfamily: glycinin

Query Match     30.6%;  Score 1018.5;  DB 2;  Length 509;
Best Local Similarity     39.1%;  Pred. No. 4.6e-57;
Matches 217;  Conservative 106;  Mismatches 157;  Indels 75;  Gaps 12;

Qy   80 DPQOQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQOREDEEKYEERM 139
Db    1 DPQRRYEOCQOEC--RQQEEROQPOCQORCLKRFEQEOQO--------- 38

Qy  140 KEGDNKRDPQOQREYEDCRRHCEQQE--PRLQYCQCORRCEQEQEQORQHGRGGDLMNPRG--- 194
Db   39 ------SQRQFEOCQOHCHQOEQORPEKKQOCVREKYQE------NPWRGERE 81

Qy  195 --GSGRYEEGEEKQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLV 251
Db   82 EEAEEEETEEGEEQOSHNPFFHFHRRSFQSRFREEHGNFRVLQREFRLS 141

Qy  252 LLEANPNAFVLPTHILDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPACTTFYLI 311
Db  142 LLEANPNTFVLPHHCDAEKIYLVTNGVVPGVVVKVPAGSTVVLA 201

Qy  312 NRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGV 371
Db  202 NQDNKEKLIIAVLHRPVNNPGQFOKFFPAGSQRPQSYLRAFSREILEPAFNTRSEQLDEL 261

Qy  372 LG------QQREGVIIRASQEQIRELTRDDSESRRWHIRRGGESSRGPYNLFNKRPLYS 424
Db  262 FGGRQSRRRQQQGQGMFRKASQEQIRALSQEAT 315

Qy  425 NKYGQAYEVKPEDYRQLQDMDVSFIANITQGSMMGPFFNTRSTKVVVVASGEADVEMAC 484
Db  316 ---SPR----EKSGE--RFAFNLLSQTPRYS 375

Qy  485 HYEQVRARLSKREAIVVLAGHPVVFVSSGN 536
Db  376 PHLPRQSSYEEEEEEEEEEEEQEEEEERRSSGQYRKIRSLSRGDIFVVPANFPVTFVASQN 435
```

```
Qy   537 ENLLLFAFG------INAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFNSQD 591
         ::|| ||| |    ||| | :|| | ||||  | |||||||  | |||||:: ||||||||
Db   436 QNLRMTGFGLYNQNINPDHNQRIFVAGKINHV-ROWDSQAKELAFGVSSRLVDEIFNSNP 494

Qy   592 ESIFFPGPROHQQQS 606
         ||| : :
Db   495 QESYFVS-RORQRAS 508
```

```
RESULT  6
A53234
globulin-1S, GLB1S - maize
C;Species: Zea mays (maize)
C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000
C;Accession: A53234; A43642
R;Belanger, F.C.; Kriz, A.L.
Genetics 129, 863-872, 1991
A;Title: Molecular basis for allelic polymorphism of the maize Globulin-1 gene.
A;Reference number: A53234; MUID:92090707
A;Accession: A53234
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-573 <BEL>
A;Note: sequence extracted from NCBI backbone (NCBIN:71280, NCBIP:71284)
A;Experimental source: inbred line Va 26
A;Title: Characterization of embryo globulins encoded by the maize Glb genes.
A;Reference number: A43642; MUID:89374022
A;Accession: A43642
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 87-100 <KRI>
C;Superfamily: glycinin

Query Match             27.1%;  Score 903;  DB 2;  Length 573;
Best Local Similarity   37.2%;  Pred. No. 1e-49;
Matches 210;  Conservative  90;  Mismatches 179;  Indels  86;  Gaps  13;
```

```
Qy    73 EEYNRQRDPQQQYEQCKCQRKCQRRETEPRHMQ-ICQQRCERRYEKEKRKQQKRYEEQQRED 131
Db   -25 EDDNHHHHGGHKSGRCVRRCEDR---PWHQRPRCLEQC-REEEREKRQERSRHEADDRSG 80

Qy   132 EEKYEERMKEGDNKRDPQQOREYEDCRRHCEQQEPRLQYQCQORCHQGRGGDLMNP 191
Db    81 EGSSE------------DERERE----- 91

Qy   192 QRGGSGRYEEGEGEEKQSD-NPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRL 250
Db    92 ------QEKEEKQKDRRPYVFDRRSFRRVVRSEQGSLRVLRPFDEVSRLLRGIRDYRV 143

Qy   251 VLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYL 310
Db   144 AVLEANPRSFVVPSHTDAHCICYVAEGGVTTIENGERRSYTIKQCHVFVAPAGAVTYL 203

Qy   311 INRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRG 370
Db   204 ANTDGRKKLVITKILHFISVPGEFQQFFPGCGRNPESFLSSFSKSIQRAAYKTSSDRLER 263

Qy   371 VLGQ--OREGVIIRASQEQIRELTRDDSE---SRRWHIRRGGESSSRGPYNLFNKRPLYSN 425
Db   264 LFGRHGQDKGIIVRATEEQTRELRHASEGGHGPHWPLPPGE-SRGPYSLLDQRPSIAN 322

Qy   426 KYGQAYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVASGEADVEMACP 485
Db   323 QHGQLYEADARSFHDLAEHDVSVSFANITAGSMSAPLYNTRSFKIAYVPNGKGYAEIVCP 382

Qy   486 HLSGRHG-------GRGGKRHEEEEVHYEQ-----VRARLSKREAIVVLAGHPVVFV 532
Db   383 HRQSGGESERERGKGRRSEEEEESSEEQEEVGQGYHTIRARLSPGTAFVVPAGHPFVAV 442
```

```
RESULT  7
B53234
viciin-like storage protein Glb1-L, embryo - maize
N;Alternate names: globulin-1L
C;Species: Zea mays (maize)
C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C;Accession: B53234; S21824
R;Belanger, F.C.; Kriz, A.L.
Genetics 129, 863-872, 1991
A;Title: Molecular basis for allelic polymorphism of the maize Globulin-1 gene.
A;Reference number: A53234; MUID:92090707
A;Accession: B53234
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-582 <BEL>
A;Cross-references: EMBL:X59083; NID:g22283; PIDN:CAA41809.1; PID:g22284
A;Experimental source: inbred line W64A6
A;Note: sequence extracted from NCBI backbone (NCBIP:71285)
C;Genetics:
A;Gene: Glb1-L
A;Introns: 167/1; 225/3; 252/3; 349/3
C;Superfamily: glycinin

Query Match             26.7%;  Score 888;  DB 2;  Length 582;
Best Local Similarity   36.7%;  Pred. No. 9.4e-49;
Matches 208;  Conservative  87;  Mismatches 181;  Indels  90;  Gaps  13;
```

```
Qy   533 SSGNENLLLFAFGINAQNHENFLAGRERNVLQQIEPQAMELAFAAASRKEVEELFNSQDE 592
Db   443 ASRDSNLQIVCFEVFADRNEKVFLAGAD-NVLQKLDRVAKALSFASKAEEVDEVLGSRRE 501

Qy   593 SIFFPGPRQ---HQQQSPRSTKQQQ 614
Db   502 KGFLPGPKESGGHEEREQEEEEREE 526
```

```
Qy    73 EEYNRQRDPQQQYEQCKCQRKCQRRETEPRHMQ-ICQQRCERRYEKEKRKQQKRYEEQQRED 131
Db    25 EDDNIHHHHGGHKSGQCVRRCEDR---PWHQRPRCLEQC-REEEREKRQERSRHEADDRSG 80

Qy   132 EEKYEERMKEGDNKRDPQQOREYEDCRRHCEQQEPRLQYQCQORCHQGRGGDLMNP 191
Db    81 EGSSED-------------EREQE- 91

Qy   192 QRGGSGRYEEGEGEEKQSD-NPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRL 250
Db    92 -----KEKQKDRRPYVFDRRSFRRVVRSEQGSLRVLRPFDEVSRLLRGIRDYRV 140

Qy   251 VLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYL 310
Db   141 AVLEANPRSFVVPSHTDAHCICVVAEGGVTTIENGERRSYTIKQGHVFVAPAGAVTYL 200

Qy   311 INRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRG 370
Db   201 ANTDGRKKLVITKILHFISVPGEFQQFFPGGGRNPESFLSSFSKSIQRAAYKTSSDRLER 260

Qy   371 VLGQ--OREGVIIRASQEQIRELTRDDSE---SRRWHIRRGGESSSRGPYNLFNKRPLYSN 425
Db   261 LFGRHGQDKGIIVRATEEQTRELRHASEGGHGPHWPLPPGE-SRGPYSLLDQRPSIAN 319

Qy   426 KYGQAYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVASGEADVEMACP 485
Db   320 QHGQLYEADARSFHDLAEHDVSVSFANITAGSMSAPLFNTRSKIAYVPNGKGYAEIVCP 379

Qy   486 HLSGRHG----GRGGKRHEEEEEVH------YEQVRARLSKREAIVVLAGHPVVFV 531
Db   380 HRQSGGESERERDKGRRSEEEEESSEEQEEAGQGYHTIRARLSPGTAFVVPAGHPFVA 439

Qy   532 VSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAASRKEVEELFNSQD 591
Db   440 VASRDSNLQIVCFEVHADRNEKVFLAGAD-NVLQKLDRVAKALSFASKAEEVDEVLGSRR 498
```

```
Qy  592  ESIFFPPGPRQ---HQQQSPRSTKQQQ 614
Db  499  EKGFLPGPEESGGHEEERDQEEEREEE 524


RESULT  8
S35221
globulin Beg1 precursor - barley
C;Species: Hordeum vulgare (barley)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C;Accession: S35221
R;Heck, G.R.; Chamberlain, A.K.; Ho, T.H.D.
  Mol. Gen. Genet. 239, 209-218, 1993
A;Title: Barley embryo globulin 1 gene, Beg1: characterization of cDNA, chromosome mappi
A;Reference number: S35221; MUID:93287988
A;Accession: S35221
A;Molecule type: mRNA
A;Residues: 1-637 <HEC>
A;Cross-references: EMBL:M64372; NID:g167003; PIDN:AAA32936.1; PID:g167004
C;Genetics:
A;Gene: Beg1
A;Map position: 4
C;Superfamily: glycinin
C;Keywords: glycoprotein
F;174-190/Product: globulin Beg1 #status predicted <MAT>

Query Match        26.0%;  Score 864.5;  DB 2;  Length 637;
Best Local Similarity   35.0%;  Pred. No. 3.2e-47;
Matches 221;  Conservative 100;  Mismatches 219;  Indels 91;  Gaps 19;

Qy   29  SKYDNQEDPQ--TECQQCQCRRCRQQESDPRQQQYCQRCRCEICEEEEYNRQRDPQQQYE  86
Db   27  ASHDEDDRGGHSLQQCVQRCQVQRCRQER--PR---YSHARCVQEC------RDDQQQH-  71

Qy   87  QCQKRCQRRETEPRHMQICQQCQVQRCRQCRCKEICEEEEYEERMKEGDNKR 146
Db   72  ------GRHEDQEEEQGRGRGWHGEGEREEEHGRGRGRHGEGEGEREEEHGRGRGRGH 119

Qy  147  DPQQREYEDCRRHCEQQEEPRLQYCQRCQRCQEQQORHGRG----GDLMNPQRGGSGRYEEG 202
Db  120  GEGEREEEEGRGRHGRHGEGE------REEERGRGRHGEGEREEEHGRGRGRHGEGEG 170

Qy  203  E--EKQSDN--PYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPN 258
Db  171  ERDEEQGDSRRPYVFGPRSFRRIIQSDHGFVRALRPFDQVSRLLRGIRDYRVAIMEVNPR 230

Qy  259  AFVLPTHLDADALLLVIGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYLLINRDNNER 318
Db  231  AFVVPGFTDADGVGVAQGEGVLTIVIENGEKRSYTVKEGDVIVAPAGSIMHLANTDGRRK 290

Qy  319  LHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVLGQQ--- 375
Db  291  LVIAKLHFISVPGKF-QFL----SVKPLLASLSKRVLRAAFKTSDERLERLFNQRQGQ 344

Qy  376  ---REGVIIRASQEQIRELTRDDSE------SRRWHIRRGGESSRGPYNLENKRPLIYSNKYGQ 429
Db  345  EKTRSVSIVRASEEQLRELRREAAEGGCOGHRWPLPFPFRCGDSRDTFNLLEQRPKIANRHGR 404

Qy  430  AYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFENTRSTKVVVVAGSGEADVEMACPHL-- 487
Db  405  LYEADARSFHALANQDVRVAVANITPGSMTAPYLNTQSFKLAVVLEGGEVOIVCPHLGR 464

Qy  488  ---SGRHGGRG------GGKRHEEEEVHYEQVRARLSKREAIVVLAG 526
Db  465  ESESEREHGKGRRREEEEDDQQQRRGSESESEEEEQQRVETVRARVSRGSSAFVVPPG 524

Qy  527  HPVFVSS--GNENLLLPAFGINAQNHHENFLAGRERNVLQQIEPQAMELAFPAASRKEVE 584
Db  525  HPVVEISSSQGSSNLQVVCFEINAERNERVWLAGR NNVIGKLGSPAQELTFGRPAREVQ 583

Qy  585  ELFNSQDESI-FPFGPRQHQQQSPRSTKQQQ 614
```

```
Db  584  EVFRAQDQDEGFVAGP---EQQSREQEQE 611


RESULT  9
FWSYBA
beta-conglycinin alpha chain precursor - soybean
C;Species: Glycine max (soybean)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C;Accession: S14681; S74124; S06714
R;Sebastiani, F.L.; Farrell, L.B.; Schuler, M.A.; Beachy, R.N.
  Plant Mol. Biol. 15, 197-201, 1990
A;Title: Complete sequence of a cDNA of alpha subunit of soybean beta-conglycinin.
A;Reference number: S14681; MUID:91355860
A;Accession: S14681
A;Molecule type: mRNA
A;Residues: 1-605 <SEB>
A;Cross-references: EMBL:X17698; NID:g18535; PIDN:CAA35691.1; PID:g18536
R;Shutov, A.D.; Kakhovskaya, I.A.; Bastrygina, A.S.; Bulmaga, V.P.; Horstmann, C.; Mu
  Eur. J. Biochem. 241, 221-228, 1996
A;Title: Limited proteolysis of beta-conglycinin alpha chain and glycinin, the 7S and 11S storage
A;Reference number: S74123; MUID:97054613
A;Accession: S74124
A;Molecule type: protein
A;Residues: 189-196,'H',198,'N',200,'X',202-203,397-408,'X',410,'X',412-417,'X',419-4
A;Experimental source: seed
C;Superfamily: glycinin
C;Keywords: glycoprotein; seed; storage protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-62/Domain: propeptide #status predicted <PRO>
F;63-605/Product: beta-conglycinin alpha chain #status predicted <MAT>
F;261,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match        25.8%;  Score 859.5;  DB 1;  Length 605;
Best Local Similarity   32.1%;  Pred. No. 6.2e-47;
Matches 203;  Conservative 133;  Mismatches 197;  Indels 99;  Gaps 18;

Qy   31  YDNQEDPQTECQQCQCRRCRQQESDPRQQQYCQRC----------KEICEEEE------EYNR  77
Db   27  YWEKENPKH--NKCLQSC-NSERDSYRNQACHARCNLLKVEEKEECEGEELPRPRPROHP  83

Qy   78  QRDPQQQYEQCQCRRCR-----RETEPRHMQICQQCQVQRCRQCRCERRY-EKEKRKQQKRYEEQQRE 130
Db   84  EREPQOPGEGEKEEDEDEQRPIPFPFPRQPQREEEHEQREQEQPRLQYCQ 143

Qy  131  DEEKYEERMKEGDNKRDPQQREYEDCRRHCEQQEEPRLQYCQRCQEQQORHGRGGDLMN 190
Db  144  EDEEDQER--QFPFPRPPHQKE---ERNEEDED--- 179

Qy  191  PQRGGSGRYEEGEEKQ-----SDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRAL 245
Db  180  -------EESEDESELRRHKNKNPFLFGSNRFETLRFKNQYGRIRVLQRFNQRSPQLONL 230

Qy  246  KNRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAG 305
Db  231  RDYRILEENSKPNTLLLENHADADYLIVTLNGTRAILSLVNNDRDSYRLQSGDALRVPSG 290

Qy  306  TTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQT 365
Db  291  TTYYVVNPDNNENLRLITLAIPVNKPGRFESFFLSSTEAQQSVLQGFSRNILEASYDTKF 350

Qy  366  ERLRGVL----------GQQR--EGVIIRASQEQIRELTRDDSESRRWHIRRGGESSRGPY 414
Db  351  EEINKVLFSREEGQQOGEQRLQESVIVIELSKEGITRALSKRAKSSSSRKTI----SSEEDKPF 406

Qy  415  NLFNKRPLIYSNKYGQAYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVA 474
Db  407  NLRSRDPIYSNKLGKFFEITPEKNPQLRDLIFLSIVDMNEGALLLPHFNSKAIVLVIN 466

Qy  475  SGEADVEMACPHLSGRHGGRGGGKRHEEEEVHYEQ-----VRARLSKREAIVVLAGHP 528
Db  467  EGDANIELV----------GLIKEQQQEEQQOEOPLEVRRKYRAELSEQDIFVIPAGYP 513
```

```
Qy  529 VVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFN 588
        || |  ||| ||| ||||||| ||  |||
Db  514 VVV--NATSNLNFFAIGINAENNQRNFLAGSQDNVISQIPSQVQELAFPGSAQAVEKLLK 571

Qy  589 SQDESIFFPGPRQHQQQSPRSTKQQQPLVSIL 620
         :||  ||| ||  ||||  :::|| |||
Db  572 NQRESYFVDAQPKKEEGNKGRK--GPLSSIL 601
```

RESULT 10
S20007
beta-conglycinin alpha chain precursor - soybean
C;Species: Glycine max (soybean)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C;Accession: S20007
R;Lelievre, J.M.; Dickinson, C.D.; Dickinson, L.A.; Nielsen, N.C.
Plant Mol. Biol. 18, 259-274, 1992
A;Title: Synthesis and assembly of soybean beta-conglycinin in vitro.
A;Reference number: S20007; MUID:92119248
A;Accession: S20007
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-605 <LEL>
A;Superfamily: glycinin

Query Match   25.8%; Score 858.5; DB 2; Length 605;
Best Local Similarity 31.9%; Pred. No. 7.1e-47;
Matches 198; Conservative 139; Mismatches 207; Indels 77; Gaps 15;

```
Qy   31 YDNQEDPQTECQQCQQCRRCRQQESDPRQQQYCQRRC-----KEICEEEE---EYNRQRDP 81
Db   27 YWEKENPKH--NKCLQSC-NSERDSYRNQACHARCNLLKVEEEECEGEIPRPRPQHP     83

Qy   82 QQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKYEERMKE 141
Db   84 EREPQPGCEKEEDEDEDEQPRPIPFPFPRQPEEEHEQEQEEEWPRKEEKRGEKGSEEEDED 143

Qy  142 GDNKRDPQQREYEDCCRRHCEQQEEPRLQYQCQRCQEQQORQHGRGGDLMNPQRGGSGRYEE 201
Db  144 EDEEQDEDRQFPFP----RPPHQEEERKQEEED--EEQQRES---------------EE 181

Qy  202 GEEKQ-----SDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEAN 256
Db  182 SEDSELRRHKNKNPFLPGSNRFETLFKNQYGRIRVLQRFNQRSPQLQNLRDYRILEENSK 241

Qy  257 PNAFVLPTHLDADAILLVIGGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNN 316
Db  242 PNTLLLPNHADADYLIVILNGTAILSLVNNDRDSYRLQSGDALRVPSGTTYYVVNPDNN 301

Qy  317 ERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVL---- 372
Db  302 ENLRLITLAIPVNKPGRFESEFLSSTEAQQSYLQGFSRNILEASYDTKFEEINKVLFSRE 361

Qy  373 -----GQQR--EGVIIRASQEQIRELTRDDSESRRWIIRGGESSRQPYNLFNKRPLYSN 425
Db  362 EGQQQGEQRLQESVIVEISKEQIRALSKRAKSSSRKYI----SSEDKPFNLRSRDPIYSN 417

Qy  426 KYGQAYEVKPEDYRQLQDMDVSVFANITQGSMMGPFFNTRSTKVVVASGEADVEMACP 485
Db  418 KLGKFFEITPEKNPQLRDLDIFLSIVDMNEGALLPHFNSKRAIVLIVINEGDANIELV- 475

Qy  486 HLSGRHGGRGGGKRHEEEEEVHYEQ-----VRARLSKREAIVVLAGHPVVFVSSGNENL 539
Db  476 ------------GIKEQQOEQQEEQPLEVRKYRAELSEQDIFVIPAGYPVV--NATSNL 522

Qy  540 LLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFNSQDESIFFPGP 599
Db  523 NFFAIGINAENNQRNFLAGSQDNVISQIPSQVQELAFPGSAQAVEKLLKNQRESYFVDAQ 582

Qy  600 RQHQQQSPRSTKQQQPLVSIL 620
Db  583 PKKEEGNKGRK--GPLSSIL 601
```

RESULT 11
B24810
beta-conglycinin alpha chain precursor - soybean
C;Species: Glycine max (soybean)
C;Date: 07-Mar-1988 #sequence_revision 30-Jun-1991 #text_change 20-Aug-1999
C;Accession: B24810; S16337
A;Molecule type: DNA
A;Residues: 1-639 <DOY>
A;Cross-references: GB:M13759; NID:g169928; PIDN:AAB01374.1; PID:g169929
A;Note: the authors translated the codon GGT for residue 352 as Glu
R;Schuler, M.A.; Schmitt, E.S.; Beachy, R.N.
Nucleic Acids Res. 10, 8225-8244, 1982
A;Title: Closely related families of genes code for the alpha and alpha' subunits of
beta-conglycinin.
A;Reference number: S16337; MUID:83143288
A;Accession: S16337
A;Molecule type: DNA
A;Residues: 361-639 <SCH>
C;Genetics:
A;Introns: 278/1; 355/3; 382/3; 481/3; 575/1
R;Doyle, J.J.; Schuler, M.A.; Godette, W.D.; Zenger, V.; Beachy, R.N.; Slightom, J.L.
J. Biol. Chem. 261, 9228-9238, 1986
A;Title: The glycosylated seed storage proteins of Glycine max and Phaseolus vulgaris
A;Reference number: A24810; MUID:86250867
C;Superfamily: glycinin
C;Keywords: seed; storage protein

Query Match   25.7%; Score 854.5; DB 2; Length 639;
Best Local Similarity 32.0%; Pred. No. 1.4e-46;
Matches 205; Conservative 141; Mismatches 215; Indels 79; Gaps 19;

```
Qy   28 WSKYDNQEDPQTECQQCQQCRRCRQQESDPRQQQYCQRRCKEI-CEEEEYN-----ROR 79
Db   28 WEKQNPSHN-----KCLRSC-NSEKDSYRNQACHARCNLLKVEEEECEGEQIPRPRPQ 80

Qy   80 DPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQRED------E 132
Db   81 HPERERQQHGEKEEDEGEQPRPFPFPRPR--QPHQEEEHEQKEEHEWPRKEERKHGGKGSE 138

Qy  133 EKYEERMKEGDNKRDPQQREYEDCRRHCEQQEEPRLQYQCQRCQEQQORQ--HGRGGDLMNP 191
Db  139 EEQDER--EHPRPHQPHQKEEE---KH-EWQHKQEKHQGKESEEEEDQDEDEQDKEESQ 192

Qy  192 QRGGSGRYEEGEEKQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLV 251
Db  193 ESEGSESQREPRRHKNKNPFHFHENSKRFQTLFKNQYGHVRVLQRFNKRSQQLQNLRDYRIL 252

Qy  252 LLEANPNAEVLPTHLDADAILLVTGGRGALKMIHRDNRESYNLECGDVIRIP------- 303
Db  253 EFNSKPNTLLLPHHADADYLIVILNGTALITLVNNDRDSYNLQSGDALRVPAGTTFYVV 312

Qy  304 ------AGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTF 352
Db  313 NPDNDENLRMIAGTTFYVVNPDNDENLRMITLAIPVNKPGRFESEFLSSTQAQQSYLQGF 372

Qy  353 SKEILEAALNTQTERLRGVL-----GQQ-----REGVIIRASQEQIRELTRDDSESRRW 401
Db  373 SKNILEASVDTKFEEINKVLFGREEGQQQGEERLQESVIVEISKKQIRELSKHAKSSSRK 432

Qy  402 HIRRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDVSVFANITQGSMMGP 461
Db  433 TI----SSEDKPFNLGSRDPIYSNKLGKLFEITQRN-PQLRLDVFLSVVDMNEGALFLP 487

Qy  462 FFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGGKRHEEEEE-VHYEQVRARLSKREA 520
Db  488 HFNSKAIVLVIVINEGEANIELV-------GIKEQQOEQQEEQPLEVRKYRAELSEQDI 539

Qy  521 IVVLAGHPVVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASR 580
Db  540 FVIPAGYPVWV--NATSDLNFFAFGINAENNQRNFLAGSKDNVISQIPSQVQELAFPRSA 597
```

```
Qy  581 KEVEELFNSQDESIFFPGPRQHQQQSPRSTKQQQPLVSIL 620
        |||||||||||||||||||||||||||  :||| |||  |
Db  598 KDIENLIKSQSESYFVDA--QPQQKEEGNKGRKGPLSSIL 635
```

RESULT 12
S00566

convicilin precursor (clone pJC4) - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 20-Aug-1999
C;Accession: S00566
R;Bown, D.; Ellis, T.H.N.; Gatehouse, J.A.
Biochem. J. 251, 717-726, 1988
A;Title: The sequence of a gene encoding convicilin from pea (Pisum sativum L.) shows th...
A;Reference number: S00566; MUID:88326208
A;Accession: S00566
A;Molecule type: DNA
A;Residues: 1-571 <BOW>
A;Cross-references: EMBL:X06398; NID:g920698; PIDN:CAA29695.1; PID:g313670
A;Note: part of this sequence, including the amino end of the mature protein, was confir...
A;Note: 237-Gln was also found
C;Genetics:
A;Gene: cvcA
A;Map position: 2
A;Introns: 221/1; 279/3; 304/3; 412/3; 507/1
A;Superfamily: glycinin
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-571/Product: convicilin #status experimental <MAT>

Query Match   25.0%; Score 833; DB 2; Length 571;
Best Local Similarity   33.2%; Pred. No. 2.7e-45;
Matches 208; Conservative 112; Mismatches 193; Indels 114; Gaps 18;

Qy   26 IDWSKYDNOEDPQTECQOCQOCQRCRQOESDPROQQYCQORCKEICEEEESNYRORDPQQQY 85
Db   25 VTYANYD--EGSETRVPGGRERGRQ 63

Qy   86 EOQCKRCQORRETEPRHMQICQQRCCERRYEKEKRKQOK---RVEEOQREDEEKYEERMKE 141
Db   64 E------KEEHE------EEKOKRYQYQREKEOKEVQPGRERWEREDEEEQVEEWRG 109

Qy 142 GDNKRDPQQOREYEDCRRHCEOQEPRLQYCQCQRCQEQOHGRGGDLMNPQRGGSGRYEE 201
Db 110 SQRREDPEER------ARLRHRER--TKRDRRH------OREGEEEERS 145

Qy 202 GEEKQSDNPYYFDERSLSTRFFTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFV 261
Db 146 SESQEHRNPFLFKSNKFLTLFENENGHIRRLQRFDKRSDLFENLQNYRLVEYRAKPHFIF 205

Qy 262 LPTHLDADAILLVIGGRGALKMIHRDNNRESYNLECGDVIRIPAGTTFYLINRDNNERLHI 321
Db 206 LPQHIDADLILVLNGRAILTVLSPNDRNSYNLERGDTFIKIPAGTTSYLVNQDDEEDLRV 265

Qy 322 AKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVL------- 372
Db 266 VDFVIPVNRPGKFEAF--GLSENKNQYLRGFSKNILEASLNTKYETIEKVLLEEQEKKPQ 323

Qy 373 ---------GQOREGVIIRASOEQIRELTLRDDSESRRWHIRRGGESSRGPYNLFNKRPL 422
Db 324 QLRDRKRTQOGEERD--AIIKVSREQIELERLKAKSSS---KKSLPSEFEPENLRSHKPE 378

Qy 423 YSNKYGQAYEVKPE--DYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVASGEADVE 481
Db 379 YSNKFGKLFEITPEKKYPQLQDIDLVSCVEINKGALMLPHYNSRAIVLLVLVNEGKGNLE 438

Qy 482 MACPHLSGRHGGRGGRGGKRHEEEEEHVYE-----QVRARLSKREAIVVLAGHPVFVFVSSG 535
Db 439 LL------------GLKNEQEREDRKERNNEVQRYEARLSPCDVVIIPAGHPVAISASS 486

Qy 536 NENLLLFAFGINAQNNIENFLAGRERNVLQOIEPQOMELAFAASRKEVEELFNSQDESIF 595
```

```
Db  487 NLNLL--GFGINAKNNQRNFLSGSGDDNVISQIENPVKELTFPGSSQEVNRLIKNQKQSHF 544
Qy  596 FPG-PRQHQQQSPRSTKQQQPLVSILD 621
Db  545 ASAEPEQKEEESQR--KRSPLSSVLD 568
```

RESULT 13
JQ1730

62K sucrose-binding protein precursor - soybean
C;Species: Glycine max (soybean)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C;Accession: JQ1730
R;Grimes, H.D.; Overvoorde, P.J.; Ripp, K.; Franceschi, V.R.; Hitz, W.D.
Plant Cell 4, 1561-1574, 1992
A;Title: A 62-kD sucrose binding protein is expressed and localized in tissues active ...
A;Reference number: JQ1730; MUID:93104680
A;Accession: JQ1730
A;Molecule type: mRNA
A;Residues: 1-524 <GRI>
A;Cross-references: GB:L06038; NID:g1431744; PIDN:AAB03894.1; PID:g170064
A;Superfamily: glycinin
A;Keywords: sugar transport
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-524/Product: 62K sucrose-binding protein #status predicted <MAT>

Query Match   23.9%; Score 794; DB 2; Length 524;
Best Local Similarity   33.4%; Pred. No. 7.1e-43;
Matches 191; Conservative 103; Mismatches 176; Indels 102; Gaps 18;

Qy   64 RCKEICEEEESNYRORDPQ-QOYEOCQKRCQRRETEPRHMQICQQRCCERRYEKEKRKQQ 121
Db   27 KCKETEVEEE------DPELVTCKHQCQQOQOYTEGCDKR---VCLQSCD-RYHRMKQERE 76

Qy 122 KRYEEOQREDEEKYEERMKEGDNKRDPQQOREYEDCRRHCEOQEPRLQYCQOCQORCQEOORQ 181
Db   77 KQIQEETRE------KKEEEESRERE 101

Qy 182 HGRGGDLMNPQRGGSGRYEEGEEKQSDNPYYFDE-RSLSTRFFTEEGHISVLENFYGRSK 240
Db 102 H------EEQDENPYIFEEDKDFETRVETEGGRIRVLKKFTESK 140

Qy 241 LLRALKNYRLVLLEANPNAFVLPHLDADAILLVIGGRGALKMIHRDNNRESYNLECGDVI 300
Db 141 LLQGTENFRLAILEARAHFVSPRHFDSEVVFNIKGRAVLGLVSESETEKITLEPGDMI 200

Qy 301 RIPAGTTFYLINRDNNERLHIA--KFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILE 358
Db 201 HIPAGTPLYIVNRDENDKLFLAMLHIPVSVSTPGKFEEFFAPGGRDPESVLSAFSWNVLQ 260

Qy 359 AALNTQTERLRGVLCQOOREGVIIRASOEQIRELTRDDSESRRWHIRRGGESSRGPYNLFN 418
Db 261 AALQTPKGRLKENVFDDQONEGSIFRISREQVRALA-PTKKSSWWPF--GGE-SKPQFNIFS 316

Qy 419 KRPLYSNKYGQAYEVKPEDYRQ--LQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVASG 476
Db 317 KRPIISNGYGRLTEVGPDDDEKSWLQRLNLMLTFTNITQGRSMSTIHYNSHATKIALVIDG 376

Qy 477 EADVEMACPHLSGRHGGRGGGKRHEEEEEHVYEQVRARLSKREAIVVLAGHPVFVFVSSGN 536
Db 377 RGHLQISCPHMS----SRSSHSKIDKSSP-SYHRISSDLKPGMWVFVVPPGHPEVTIASNK 431

Qy 537 ENLLLFAFGINAQNNHENFLAGRERNVLQOIEPQOAMELAF 581
Db 432 ENLMICFEVNARDNKKFTFAGRD-NIVSSLDNVAKELAFNYPSEMVNGVFLLQRFLERK 490

Qy 582 EVEELFN----SQDESIFFPGPRQHQQQSPRS 609
Db 491 LIGRLYHLPHKDRKESFFFELPREERGRRA 522
```

RESULT 14

```
T06459
62k sucrose-binding protein homolog - garden pea
N;Alternate names: p54 protein
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C;Accession: T06459
R;Rodrigo II, M.I.
submitted to the EMBL Data Library, February 1997
R;Reference number: Z15693
A;Accession: T06459
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-483 <ROD>
A;Cross-references: EMBL:Y11207; PIDN:CAA72090.1
A;Experimental source: var. Lincoln
C;Genetics:
A;Note: p54
C;Superfamily: glycinin

Query Match   23.38;    Score 775.5;   DB 2;   Length 483;
Best Local Similarity 30.4%;   Pred. No. 9.5e-42;
Matches 170;  Conservative 112;  Mismatches 163;  Indels 115;  Gaps 12;

Qy    50 QQESDPRQQQYCQRRCKEICEEEEENRQRDPQQQYEQCQKRCQRRETEPRIMQICQQRC 109
            :|:||              |:|||||  ||:||:||  |  :||||  :||
Db    28 RKEKDPE-------------DMQRQYDEEDKR-----------------ICMERC 60

Qy   110 ERRYEKEKRKQQKRYEQQQREDEEKYEERMKEGDNKRDPQQREYEDCRRHCQQQEPRLQY 169
            :|:|| ||  ||:::|| |:||||:|||:||
Db    61 D-DYIKKKQERQKHKEHEEEEQEQEED------------- 87

Qy   170 QCQRRCEQRQHGRGGDLMNPQRGGSGRYEEGEEEKQSDNPYYFDERSLSTRFRTEEGHI 229
Db    88 -------------------ENPYVFEDNDFETKIDTKDGRV 109

Qy   230 SVLENFYGRSKLLRALKNYRLVLLEEANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNR 289
Db   110 LILNKFNEKSKLLKNIENYGLAVLEIKANAFLSPHHYDSEAILFNIKGRGIIGLVAEDRT 169

Qy   290 ESYNLEGGDVIRIPAGTTFYLLINRDNNERLHIAKF---LQTISTPGQYKEFFPAGGQNPE 346
Db   170 ERFNLEEGDIMRVPAGTPMYLVNRDENEKLYIAAFIHMPPSSGSAPVNLEPPFESAGRKPE 229

Qy   347 PYLSTFSKEILEAALNTQTERLRGVLGQQREGVIIRASQEIRELTRDDSESRRWHIRRG 406
Db   230 SVLNTFSSKVLQAALKSSKGELETVLDEQKGKRIFKIEKEDVRGLAPKKS---LWPF--- 283

Qy   407 GESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQ-LQDMDVSVFIANITQGSMMGPFENT 465
Db   284 GGPFKSPFNIFSNNPAFSNKFCSLFEVGPSQEKSGLEGLNLMLTLANITKGSMSTIHYNT 343

Qy   466 RSTKVVVVASGEADVEMACPHLSGRHGQRGGGSKRHEEEEEVHYEQVRARLSKREAIVVLA 525
Db   344 NANKIALVIDGEGELEMACPIHMPS-----SSSNSRQKKSSISYHNINAKLRPGVMFVVPA 398

Qy   526 GHPVVFVSSGNENLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEE 585
Db   399 GHPFVNIASKKNLIVVCFEVNAQRNNKLALAGK-KNIVSALDKAAKEVAFDIAAEKVDE 457

Qy   586 LFNSQDESIFFPGPRQHQQQ 605
            :|| |
Db   458 VFERKEE-FFPYDNEERKE 476


RESULT 15
S00567
vicilin precursor (clone pDUB9) - garden pea (fragment)
C;Species: Pisum sativum (garden pea)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Aug-1999
C;Accession: S00567
R;Watson, M.D.; Lambert, N; Delauney, A.; Yarwood, J.N.; Croy, R.R.D.; Gatehouse, J.A.;
Biochem. J. 251, 857-864, 1988
```

```
A;Title: Isolation and expression of a pea vicilin cDNA in the yeast Saccharomyces ce
A;Reference number: S00567; MUID:88326226
A;Accession: S00567
A;Molecule type: mRNA
A;Residues: 1-433 <MAT>
A;Cross-references: EMBL:Y00722; NID:g20915; PIDN:CAA68708.1; PID:g758248
C;Superfamily: glycinin
F;1-2/Domain: signal sequence (fragment) #status predicted <SIG>
F;3-433/Product: vicilin #status predicted <MAT>

Query Match   22.8%;    Score 759.5;   DB 2;   Length 433;
Best Local Similarity 37.4%;   Pred. No. 8.5e-41;
Matches 161;  Conservative 81;  Mismatches 154;  Indels 35;  Gaps 5;

Qy   209 NPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEEANPNAFVLPTHLDA 268
            |:|  |            |||  ||:||:||||  :|||:|| |||:||||:| |||||| ||
Db     8 NPFIFKSNKFQTLFENENGHIRLLQKFDQRSKIFENLQNYRLLEYKSKPHTIFLPQHTDA 67

Qy   269 DAILLVIGGRGALKMIHRDNRESYNLEGGDVIRIPAGTTFYLLINRDNNERLHIAKFLQTI 328
            |:|| ||   | :|::|:||:| :|||| || ::|||||||:::|||::|:|:||:|:||
Db    68 DYILVVLSGKAILTVLKPDDRNSFNLERGDTIKLPAGTTAYLVNRDDNEELRVLDLAIPV 127

Qy   329 STPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVL-------------- 372
            : |||::|  | |||| ||:||||| |:|||:||   ||::::||
Db   128 NRPGQLQSFLLSGNQNQNYLSGFSKNILEASNTDYEIEKVLLEEHEKETQHRRSLKD 187

Qy   373 --GQQREGVIIRASQEIREL?RDDSESRRWHIRRGGESSRGPYNLFNKRPLYSNKYGQ 429
              ||||:::|::|::|:|||  |      |::|:|||:||||||:|||:||:|||:|
Db   188 KRQQSQEENVIVKLSRGQIEELSKNAKSTS----KKSVSSESEPFNLRSRGPIYSNEGK 243

Qy   430 AYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFENTRSTKVVVVASGEADVEMACPHLSG 489
            |:|:|||| ||||:|:|:|  ::|:|||:||||::||  |:|::||:|  |:|||:||| :
Db   244 FFEITPEKNPQLQDLDIFVNSVEIKEGSLLLPHYNSRAIVIVTVNEGKGDFELVGQRNEN 303

Qy   490 RHGGRGGGKRHEEEEEVHYEQV----------RARLSKREAIVVLAGHPVVFVSSGNENLLF 542
            ::  :|  |:::|||::|:|:             |:|::||:| :||:|||::|:|:||:|
Db   304 QQEQR---KEDDEEEQGEEINQVQNYKAKLSSGDVPVIPAGHPVALKASSNLDLL-- 358

Qy   543 AFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFNSQDESIFFPGPRQH 602
            ||||||:||:|||:||:|:|||| |  |||||||| :|:||:|: |::: ||||||||:|
Db   359 GFGINAENNQRNFLAGDEDNVISQVQRPVKELAFPGSAQEVDRILENQKQSHFADAQPQQ 418

Qy   603 QQQSPRSTKQQ 613
            ::|: |:  :
Db   419 RERGSRETFDR 429
```

```
Search completed: March  1, 2001, 15:52:36
Job time: 560 sec
```